Strathprints Institutional Repository

Lalmas, M. and Ruthven, I. (1999) *A framework for investigating the interaction in information retrieval.* In: Information Modelling and Knowledge Bases XI, 1999-05-24 - 1999-05-28, Iwate, Japan.

20pt 20pt

A framework for investigating the interaction in information retrieval

Mounia Lalmas*
Department of Computer Science
Queen Mary and Westfield College, University of London
Ian Ruthven
Department of Computing Science
University of Glasgow

Abstract. To increase retrieval effectiveness, information retrieval systems must offer better supports to users in their information seeking activities. To achieve this, one major concern is to obtain a better understanding of the nature of the interaction between a user and an information retrieval system. For this, we need a means to analyse the interaction in information retrieval, so as to compare the interaction processes within and across information retrieval systems.

We present a framework for investigating the interaction between users and information retrieval systems. The framework is based on channel theory, a theory of information and its flow, which provides an explicit ontology that can be used to represent any aspect of the interaction process. The developed framework allows for the investigation of the interaction in information retrieval at the desired level of abstraction.

We use the framework to investigate the interaction in relevance feedback and standard web search.

1 Introduction

Modern *information retrieval* (IR) is an inherently **interactive** process and users can expect to engage in a variety of tasks and techniques in the course of an information seeking session. For instance, what users learn about a system (e.g., how to search in a particular system, how to express an information need, how to perform relevance feedback) is usually seen as the result of their interaction with the system.

In recent years, there has been a growing awareness of the importance and diversity of the user's information seeking behaviour and interaction styles. This is because the amount of available information is currently growing at an incredible rate (a particular example of this is the Internet), and more people with different backgrounds are confronted with the problem of finding relevant information. To provide effective retrieval, it is therefore becoming mandatory to offer better support to users in their information seeking activities. This requires a better understanding of the interaction process in IR. This is for example one of the aims of the TREC conference, with the introduction of the "interactive track" [13].

A number of researchers have attempted to classify various aspects of information seeking activities and proposed theories (for example, [10, 5, 18, 15, 3]). However these theories do not provide a formalism in which the interaction can be formally expressed and then studied and compared. This means that we cannot *reason* about the interaction occurring between a user and an IR system. As a result, we cannot predict which system will better support which user. To obtain a better understanding of the nature of interaction in IR, we need a **formal framework** to investigate the interaction between a user and an IR system.

We use **channel theory** [1, 8, 2] to construct a formal framework for representing and studying the interaction in IR. In [?, ?], we used situation theory, upon which channel theory is based, to construct

*This work was carried out when the author was at Informatik VI, University of Dortmund, Germany.

logic-based IR models and meta-models [?]. At that time, we were concerned with capturing the information content aspect in IR. Later, in [22, 14, 19], we demonstrated the expressiveness and the appropriateness of situation theory and channel theory in capturing many features present in IR. The features can be standard, such as document and query representations, matching function (information content aspect); or they can be features that have not traditionally been considered to be part of the IR system, such as user interaction, hypertext and multimedia data[1]. This study constitutes a first major step towards the development of an expressive framework for the modelling of IR. A major benefit of having such a framework is that a *general and uniform* framework is acquired. This allows for the theoretical comparison of IR systems not only based on their effectiveness, but also with respect to their characteristics (e.g., precision vs. recall oriented systems, navigation vs. querying, usability, etc). In the present work, we extend on one aspect: the investigation of the interaction in IR.

Channel theory provides a “science” for the representation and manipulation of *information and its flow*. The latter characterises the notion of *information containment* and can generally be defined as the information an object contains about another. The interaction process in IR involves steps, the objects, (e.g., user browsing hypermedia documents, user reading documents, user selecting relevant documents for relevance feedback), AND the information available in a step usually contains information about the next step. Furthermore, because it follows a mathematical approach, channel theory provides an explicit ontology (and not merely a formal device as in the possible-worlds approaches) that can be used to capture both specific and general aspects of the interaction process. As a result, a channel-based framework allows for the investigation of the interaction in information retrieval at the **desired level of abstraction**.

The paper is organised as follows. In Section 2, we give an overview of interactive activities in IR, explaining at the same time, why we need to model interaction. In Section 3, we describe some theories that were developed for representing the interaction, and discuss their limitations. In Section 4, we present our framework based on channel theory. In Section 5, we apply our model to two cases: query reformulation and standard web search. We finish with a conclusion and some thoughts for future work.

2 Areas of information retrieval interaction

The areas of interaction can be roughly grouped into two types: those that occur across sessions and those that occur within sessions.

In the first group we can include (1) learning about a particular information source, (2) learning what facilities a particular system offers, and (3) the development of information seeking strategies. Techniques to support the first activity tend to operate at the interface level with techniques such as visualisation, semantic maps, and clustering, all designed to display relationships between concepts, documents, terms, etc within a particular database. The other two activities are strongly interlinked. As described in [12, 11] and others, people come to IR systems with existing information seeking strategies but also develop strategies for using particular IR systems. The success of searching and the users satisfaction with a system do not necessarily depend on what interactive features a system offers or on how it encourages the users to use these features but on how well the system supports the users strategies and how well it leads the user to understanding how it operates [6]. These factors will also be influenced by how experienced the users are at searching and how experienced they are with a particular system.

Within a search session, the two main areas of interaction can be grouped into query formulation and post-query interaction (of which the basic operation is relevance feedback). These areas are not mutually exclusive since some systems do not force the user to explicitly formulate a query and some relevance feedback techniques may involve manual query reformulation. Initial query formulation is important since the choice of terms a user enters may determine the success of the whole search. How the user enters query terms, the use of controlled vocabularies, the use of complex query specifications [7] are all aspects of the interaction that can affect the user’s understanding of the system and the effectiveness of the retrieval.

Relevance feedback, in general, tries to bring the query closer to the user’s actual information need by using the user’s relevance assessments (those documents marked explicitly relevant by the user). This may be done automatically by re-weighting terms and/or adding new terms from relevant docu-

[1]Although these features have existed for some time and their impact on retrieval has been studied, formally and informally, they have not ben treated as integral to the study and representation of IR systems

ments to the query (query expansion). These terms may also be derived using semantic relationships [24], co-occurrence information [23] or frequency information [21]. Query expansion can be performed automatically (the system decides which terms to add) or interactively (the user decides which terms to add, where the terms come either from a system generated list or from a knowledge base such as a thesaurus).

In [20] and [4], it is argued for an increased user control over interaction, but it was also demonstrated that in practical cases this does not necessarily translate into improved retrieval effectiveness. For instance, in practice, we have good principles for how to apply relevance feedback, but the actual performance varies. Automatic relevance feedback effectiveness is very variable but overall fairly effective, whereas the performance of interactive relevance feedback is more consistent but not as effective [20]. The difficulty is that although we can derive principles for predicting when relevance feedback is likely to improve or degrade retrieval effectiveness, these principles tend to be very general. We have no means of comparing and predicting the effect of the interaction at a more precise level. Also, the principles tend to apply to the search as a whole, and not to individual stages of the search. For example, automatic query expansion tends to produce better results on short queries, whereas interactive query expansion works better in data rich environments.

Thus we need a means of representing the interaction within and across IR systems at the desired level of abstraction. This will allow us to compare the interactive processes in IR. This comparison also affects the information seeking strategies as the detail of the interaction over the course of a session allows us to examine *how* the user is searching for information. The comparison would allow an evaluation, thus explaining effectiveness or non-effectiveness of some interactive processes offered by some systems. We therefore need a **meta-theory** for representing the information seeking activity in IR.

3 Information seeking theories

A number of researchers have attempted to classify various aspects of information seeking. Their contribution is not simply to explain what users do with particular systems but to describe what they do in the information seeking process as a whole. We shall study three theories of information seeking which look at different modelling aspects: the range of information seeking behaviours, the complexity of interaction and behaviour, and how to support information seeking in IR systems.

3.1 Range of information seeking behaviours

In an attempt to understand how people search, several researchers have produced categorisations of information seeking. We look at three of them.

The first categorisation comes out of a long running exploration of user's interaction. Belkin et al [5] have defined a general, high-level classification of information seeking strategies (ISS) that users can employ within the course of an information seeking session. This has led them to propose a single framework for interactive IR systems based on the hypothesis that users not only engage in multiple ISSs across searches but also within a search. They classify ISSs based on their value on each of four dimensions. These dimensions are presented as discrete binary values but Belkin et al acknowledge that they can be regarded as a continuum. Each dimension represents an aspect of information seeking:

1. method of searching:- scanning vs searching. The method of searching describes what the user is using the system for; searching implies that a user is looking for a particular piece of information, whereas scanning implies that a user is browsing or scanning for an item of interest.

2. mode of retrieval :- recognition vs specification. The mode of retrieval describes how the user is searching, either by specification (e.g., a query or some other description of a need) or by recognition, that is recognising an item as being relevant rather than describing what makes it relevant.

3. goal of retrieval :- learning vs selection. The goal of retrieval specifies the purpose of retrieval at that particular point in the search. A user may want to learn more about the system, or collection, or may be selecting items for retrieval or feedback.

4. resource considered :- information vs meta-information. The resource considered describes what the user is using to interact with, either meta-information such as a thesaurus or index, or the information (documents) itself.

The combination of these dimensions give 16 possible ISSs. Moving from one ISS to another is an example of the movement from one information seeking strategy to another. ISSs are viewed, at least conceptually as, discrete activities (*steps*) rather than the adaptive movement (*transition*) from one step to another. How one moves from one ISS to another is dictated by the particular implementation of the system, which must determine how to use the information the user is passing from one ISS to another. In Belkin et al's theory, the ISSs describe what the user wants but not the transitions between ISSs. Describing these transitions is important in a theory for interactive IR. Nevertheless they remain a useful list of searching behaviours that characterise the interaction between users and IR systems.

Ellis [9] also presents a categorisation based on the behaviour of searchers and their relation to information seeking design. The study in [9], based on the information seeking patterns of academic social scientists, was extended in [10] with a study of academic physics researchers. The result was a categorisation of eight activities, each subsuming a variety of actual behaviours, which in combination can represent any individual's information seeking pattern. The eight activities: starting, chaining, browsing, differentiating, monitoring, extracting, verifying and ending, range from specific information seeking (e.g., chaining) to more general information management techniques (e.g., monitoring). Ellis' theory provides a means of describing the individuals seeking patterns but does not seek to predict either the *order* or *combination* of stages. Unlike Belkin, Ellis' categorisation concentrates not on what users are doing but on what task they are trying to achieve with the system.

Kuhlthau [17, 18] presents an alternative categorisation of steps[2]: initiation, selection, exploration, formulation, collection, and presentation. This theory incorporates affective, cognitive, and physical aspects. For example, in an exploration phase, uncertainty and confusion may increase (affective), thoughts centre on gaining a focus (cognitive) and physical activities are guided towards locating relevant information and relating new information to existing information. This is a theory of an idealised search in which Kuhlthau promotes *uncertainty* as the primary factor in characterising the move from one step to another and the particular choice of activity. The overall search is seen as moving from initial "vague thoughts, anxious feelings, and exploratory actions" to understanding, "clear thoughts, confident feeling and documentary actions".

Each of these studies provides a different slant on information seeking. Belkin et al theory ways of finding information in a search, Ellis et al theory specific activities including complex strategies, and Kuhlthau includes factors that motivate why the user chooses certain strategies at certain points in the search. These are not different ways of representing the same thing: each brings new aspects of the search process to light and each requires to be supported in an interactive system. Therefore, none of the frameworks can be used to investigate interaction processes with the aim to compare them.

3.2 Complexity of search activity

The theories presented in the previous section describe aspects of searching that are independent of the system being used. However, these generic activities or aspects must ultimately be mapped onto a series of actions performed by the user on the system. What these theoriess demonstrate is that user strategies can be relatively abstract or undetermined, potentially creating a gulf between what users want to achieve and how the system supports them achieving their goal. This ease or difficulty of this translation - from conceptual goal to a series of user actions - can influence how the user uses the system.

This aspect of information seeking was examined by Bates [3] who looked at the granularities of behaviour that occur within, across and outside of searches. This analysis allows her to speculate on how IR could support interactive IR with the rather bleak observation that most, current IR systems only support very limited interaction with very little support for more complex information seeking patterns. She classifies levels of interaction into four kinds:

- Level 1: move. This corresponds to a single action or thought and is the basic unit of interaction corresponding to a single action such as marking a document relevant or selecting a query term. She argues that most IR interaction is, at present at this level.

- Level 2: tactic. One or a handful of moves such as MONITORING TACTICS (e.g., record track of trails followed) or SEARCH FORMULATION TACTICS (e.g., try again under a different spelling). Some IR interaction is moving towards this but often although the action may be supported (e.g., modifying the query to try an alternate word) the user is forced to do each step manually, and there is little support for comparing actions.

[2]Unlike Ellis' theory, Kuhlthau does assert an ordering of steps.

- Level 3: stratagem. This is a larger, more complex set of moves, composed of multiple tactics all designed to exploit the file structure of a particular search domain thought to contain all desired information, for example, locating an area of interest then browsing, or starting with a citation and following all works that have cited it. What she calls stratagems could correspond to IR systems - different approaches to collecting information using a particular interaction style. The whole operation from querying to presenting information takes place in one go but of course most IR systems are supporting one type of search for all domains and (notionally at least) all types of search.

- Level 4: strategy. A plan, may contain moves, tactics and/or stratagems for an entire information search. This is the top level of interaction, the most complex and also the most difficult to adapt to IR because users adapt to the information being presented by the system.

The properties of each level are "emergent". That is each, conceptually and practically, is more than the sum of its subcomponents. So, although ultimately every level is composed of a series of moves, it is not the conjunction of particular moves nor the complexity of these moves that differentiate between a stratagem or a tactic. Rather it is what the user is trying to achieve by the interaction that separates them.

Bates' approach connects well to empirical evidence about what techniques users employ when using an IR system but is not intended to serve as a investigating appraoch. The strength of the approach is that it distinguishes between the complexity of user's information seeking tasks. Its weakness for investigating interactive IR in the way that we are interested in doing is that it is not intended to investigate individual searches.

3.3 Supporting information seeking

A third approach to characterise behaviour for IR systems is a functional approach taken by Ingwersen in his work on the Mediator Model [15]. The Model, based on theoretical and empirical research, aims to support the construction of user models in particular but can be generalised to other approaches. The focus of Mediator is not the modelling of user or system but that of an intelligent intermediary to support searching.

Mediator is composed of thirteen units or functions, each corresponding to a particular knowledge source or mechanism. We shall not describe these in detail here but they can be classified into 3 types: those that contain knowledge about work tasks, system or user; those that generate knowledge about the database, information need, user; and those that select an IR strategy, examine user response and interact with the user and other components of the system.

Such a rich contextual operating system for interaction relies on adequately specifying *a priori* the range of behaviour an intermediary should support. It also raises the issue, outlined by Bates [3], of where the control over the interaction should lie: with the user or with the system?

Although this approach does not force the user to perform specific actions, it does promote a very specific information seeking framework. Therefore, it cannot serve as a representation language in which different approaches to supporting searches can be compared and contrasted.

3.4 Conclusion

We have seen a number of theories for interactive IR systems. These theoriess tend to fall into one of two types: *descriptive* or *prescriptive*. The descriptive theoriess, e.g., [10, 18], describe what stages or processes are commonly seen in interaction. They attempt to categorise or structure the phenomena associated with information seeking. The prescriptive theoriess, e.g., [5, 15], outline what activities IR systems should support and, to varying degrees, how they should support them. These two approaches can be seen as complementary: one describing the nature of information seeking and the other detailing how this should be practically supported.

However, these theoriess do not provide a formalism in which the interaction can be formally expressed and then studied and compared. This means that we cannot *reason* about the interaction occurring between users and systems. The descriptive theoriess can be used to characterise information seeking patterns but they do not allow us any predictive power: they cannot predict which systems will better support which users. Also they have a strong element of *interpretation*; since user's actions must be mapped onto a model, they do not allow us to analyse the search formally. The prescriptive theories on the other hand can lack generality because user's actions correlate with specific system features. **Consequently we need a meta-theory to investigate the interaction between IR systems and users.**

4 A meta-theory for investigating interaction

We present a formal framework based on **channel theory** [1, 8, 2] to represented and study the interaction in IR. Based on the approaches described in the previous section, we identify two aspects to be captured: the *steps* of the interaction, and the *transition* between steps. A third aspect is the *reasoning* about the interaction. We refer to them as, respectively, the **static aspect**, the **dynamic aspect** and the **reasoning aspect**.

4.1 Static aspect of the interaction

Representing the static aspect of the interaction consists of repersenting the steps involved during a user's interaction with the IR system. To represent a step, we must symbolise the step itself and the knowledge associated with that step, that is the *information* true at that step.

We consider two forms of information: "A property R holds/does not hold for the objects $a_1, \ldots, a_n$". The two forms are modelled by two *infons*, respectively: $\ll R, a_1, \ldots, a_n; 1 \gg$ and $\ll R, a_1, \ldots, a_n; 0 \gg$. For example: the infon $\ll Browsing, docA, user3; 0 \gg$ represents the item of information that a user, user3, is not browsing a document, docA; the infon $\ll Indexing, query3, wine, 0.56; 1 \gg$ represents the item of information that wine is a query term for query3 and its weight is 0.56.

Nothing is said so far about the truth or falsity of an infon. An infon is true or false with respect to a context, referred to as a *situation*. Let σ be an infon and s a situation. The fact that s contains or "make true" σ is modelled by $s \models \sigma$. We say that s "supports" σ.

A situation represents one step in the interaction, and the infons it supports constitute the information true at that step. For example, during relevance feedback, the step where the reformulated query is submitted to the IR system corresponds to a situation. The items of information (infons) supported by a situation are the terms and the weights forming the query used at that step.

Consider the two infons $\ll Searching, userA, 1pm; 1 \gg$ and $\ll Browsing, userA, 3pm; 1 \gg$. The infons have the common information that a user, userA, is doing some action. What differs is the action itself and the time of its occurrence. Such uniformities among infons are represented by *types*. For example, the type abstracting among the two previous infons is $action\_userA = [\dot{s} | \dot{s} \models \ll \dot{a}, userA, \dot{t}; 1 \gg]$: the type of any situation about an action (represented by the parameter $\dot{a}$) and a time (represented by the parameter $\dot{t}$) performed by userA. If the situation s is one of them, this is written $s \models action\_userA$[3].

Types are used to represent standard concepts, activities or knowledge describing the steps in the interaction between a user and an IR system. For example, the type of situations where a user is browsing is represented by $browsing = [\dot{s} | \dot{s} \models \ll Browsing, \dot{u}; 1 \gg]$; the type of situations where a user is doing no action is repersented by $no\_action = [\dot{s} | \dot{s} \models \ll \dot{a}, \dot{u}; 0 \gg]$.

4.2 Dynamic aspect of the interaction

Representing the dynamic aspect of the interaction consists of representing (1) the (physical) transitions between steps, (2) the nature of the transitions, that is the information flowing from one step to another, and (3) whether the transitions are uncertain.

4.2.1 Transitions between steps

Since steps are represented as situations, the physical transition between steps is the passage of one situation to another. This is represented by *channels*, which are relations between situations. Let c be a channel that connects situation s_1 to situation s_2. This is written $s_1 \overset{c}{\longmapsto} s_2$. It expresses that there is a *flow of information* from situation s_1 to situation s_2: the information available at s_1, the *source* situation, contains information about the situation s_2, the *target* situation, and this with respect to channel c.

Consider the case of a user browsing a document D. Let this step be represented by situation d. Suppose that the document D is linked to a document D'. If the user decides to browse D' (by clicking

[3]In [8], a detailed description of infons and situations, together with a set of rules that ensure proper instantiating (called anchoring) of parameters, is given.

on a hypertext link in document D), then another step arises, one when the user is browsing D'. Let this step be represented by situation d'. The information supported by d contains information about d' because information flows from d to d' (due to the existence of the hypertext link). The relation between d and d' is represented by a channel.

Let s be a situation representing the step, in relevance feedback, where the user is selecting among the retrieved documents those which are relevant. The reformulation of the query leads to another step, represented by a situation s'. Information flows from situation s to situation s', since the reformulated query is based on the original one. This is represnted by having a channel relating s to s'.

4.2.2 Nature of the transitions

The transitions between steps have different natures. For example, in a manual query expansion process, terms synonyms to those used in a query (a step) can be used to expand the query (the next step). The nature of transition is based on synonymy information (extracted for instance from a thesaurus).

Let c be the channel linking the situation s_1 to the situation s_2. The nature of the transition is expressed by a *constraint* $\varphi \rightarrow \psi$, where φ and ψ are types, if whenever $s_1 \models \varphi$ then $s_2 \models \psi$. It is said that $s_1 \models \varphi$ carries the information that $s_2 \models \psi$, and that the channel c is of type or support $\varphi \rightarrow \psi$. This is written $c \models \varphi \rightarrow \psi$. The whole process (the channel and its nature, and the linked situations and their types) is represented as follows: $s_1 \models \varphi \stackrel{c}{\longmapsto} s_2 \models \psi$.

4.2.3 Uncertain transitions

The transitions between steps may be uncertain. For example, suppose that a user is in a browsing situation s. Depending on the information available at that step, the user can either continue browsing, or query the system directly. The transition from the step represented by the situation s and the next step represented by a situation s' is uncertain because it may not be the case that the user will end up in that next step. The uncertainty depends on the information available in situation s (e.g., the user intention, the user satisfaction).

Since the nature of a transition is represented by constraints, we must differentiate between constraints that always hold, the *unconditional* constraints, and those that do not, the *conditional* constraints. With the former, the transitions always occur, whereas, with the latter, they may not occur. The uncertainty is captured by *background conditions*. A conditional constraint is written $\phi \rightarrow \varphi | B$, which highlights the fact that the constraint $\phi \rightarrow \varphi$ holds if the background conditions captured within B are *supported by the source situation*.

Consider the example of the manual query expansion process discussed above. The terms used in the query may be ambiguous. Hence a transition based on synonymy information is uncertain, since the terms synonymous to those used in the query (which will be used to expand the query) depend on the sense of the terms in the original query. The constraints here relate terms to their synonyms, and the background conditions capture the term senses.

4.3 Reasoning aspect of the interaction

The purpose of a meta-theory of the interaction in IR is not just to describe the interaction, but also to reason about it. Channel theory defines six basic operations that allow for reasoning about information and its flows. The operations applied in our context enable us to modelcapture the followings: (1) a transition cannot occur; (2) no change occurs; (3) transitions can happen in sequence; (4) several transitions may occur between two steps; (5) transitions can be invertible; and (6) a transition is better than another (the "goodness" criterion depends on the context). These six cases are not exhaustive, but cover the most common users' information seeking behaviour in IR.

4.3.1 A transition cannot occur

Not all pairs of situations are linked. The fact that a transition from a situation s to a situation s' cannot occur is represented by the *null channel* denoted 0. It means that there is no flow or there is an inconsistent flow from s to s'.

4.3.2 No change occurs

A transition can relate a step to itself. For example, in a hypertext, a document is linked to itself. When following the link, the user remains in the document, so there is no change with respect to the actual step. This is represented by the *unity channel* denoted 1. This channel links a situation to itself: $s \xmapsto{1} s$. If the nature of the channel is given by the constraint $\varphi_1 \rightarrow \varphi_2$, that is, $1 \models \varphi_1 \rightarrow \varphi_2$, then: if $s \models \varphi_1$, we also have $s \models \varphi_2$.

4.3.3 Sequential transitions

Transitions can happen in sequence since sequential processes occur during the interaction between a user and an IR system. A sequence of transitions is represented by the *sequential combination of channels*. Two channels c_1 and c_2 can be combined sequentially, denoted $c_1; c_2$, if the target situation of c_1 is a source situation for c_2. $c_1; c_2$ constitutes a channel. If no source situation for c_2 is a target situation for c_1, then $c_1; c_2$ corresponds to the null channel 0. An example of this would be to forbid the following sequence of user actions: first browsing a system and then querying it.

Let $s_1 \models \psi_1 \xmapsto{c_1} s_2 \models \psi_2$ and $s_2 \models \psi_2 \xmapsto{c_2} s_3 \models \psi_3$. The nature of the transitions represented by c_1 and c_2 is given by the constraints $\psi_1 \rightarrow \psi_2$ and $\psi_2 \rightarrow \psi_3$, respectively. We have then $s_1 \xmapsto{c_1;c_2} s_3$ and the nature of the channel $c_1; c_2$ is given by the constraint $\psi_1 \rightarrow \psi_3$. That is, we have $s_1 \models \psi_1 \xmapsto{c_1;c_2} s_3 \models \psi_3$

4.3.4 Parallel transitions

There may be parallel transitions between two steps. For example, in a query expansion process, a user can modify his/her query, and the system can determine the new terms to be added to the query. Let us consider the step, represented by a situation s_1, where documents have been retrieved for the initial query, and the step, represented by the situation s_2, where the list of terms to be added or removed to the query has been fixed. Each strategy can be represented as a channel (or a sequence of channels), respectively c_1 and c_2 where $s_1 \xmapsto{c_1} s_2$ and $s_1 \xmapsto{c_2} s_2$. We have two parallel transitions.

The parallel transitions are formally expressed as the *parallel combination* of channels. Two channels c_1 and c_2 can be combined in parallel, denoted $c_1 \| c_2$, if for any situation s_1 and s_2 such that $s_1 \xmapsto{c_1} s_2$, we also have $s_1 \xmapsto{c_2} s_2$, and vice versa. $c_1 \| c_2$ constitutes a channel relating s_1 to s_2. Let $s_1 \models \psi_1 \xmapsto{c_1} s_2 \models \psi_2$ and $s_1 \models \psi_3 \xmapsto{c_2} s_2 \models \psi_4$. The channel $c_1 \| c_2$ supports the constraint $(\psi_1 \wedge \psi_3) \rightarrow (\psi_2 \wedge \psi_4)$. That is, we have $s_1 \models (\psi_1 \wedge \psi_3) \xmapsto{c_1 \| c_2} s_2 \models (\psi_2 \wedge \psi_4)$.

4.3.5 Inverse transitions

In a hypertext system, a user browsing through one document to another may decide to retract and return to the document previously visited. Going from a document D to a document D' can be represented by a channel, and going from document D' to document D is represented by the *inverse* channel.

For a channel c linking the situation s to the situation s', the inverse channel denoted c^{-1} links s' to s. It should be noted that it is not because $c \models \phi \rightarrow \psi$ that we have $c^{-1} \models \psi \rightarrow \phi$. This is one main difference between channels and constraints. Channels are invertible, but constraints are not automatically. Also, it can happen that the inverse channel corresponds to the null channel. This would mean that a transition occurs between two steps, but there is no reverse transition.

4.3.6 Comparing transitions

To reason about the interaction between a user and an IR system, we need a means to compare transitions. Since the ultimate aim of using an IR system is to find relevant information, one approach would be to compare the ease of finding relevant information.

This comparison approach is represented by the notion of *refinement*. A channel c_1 is a refinement of a channel c_2, written $c_1 \preceq c_2$ if and only if $c_1 = c_1 \| c_2$. In other words, c_1 is a refinement of c_2, if any information one obtains using c_1 and c_2 in parallel could be obtained using c_1 alone.

Suppose that channels represnt navigational paths in a browsing system. Let c_1 be a channel representing a path between two documents that is shorter than a second path, represented by a channel c_2. This can be proved by showing that $c_1 \preceq c_2$.

5 Application

In the previous section, we described how and why channel theory can be used as a meta-theory to represent and study the interaction in an IR system. The proposed framework can be used to formally investigate the interaction in an IR system in real search environments. The investigation will provide results regarding how users interact both within and across IR systems using identical or different information-seeking techniques (e.g., relevance feedback, browsing, interactive query expansion). These results would provide crucial input for developing more usable and hence more effective IR systems.

In this section, we show how the proposed framework can be used to formally investigate interaction in two cases: query reformulation and web searching.

5.1 *Query reformulation*

Query reformulation is the method by which queries are modified in order to bring the query closer to the user's actual information need as indicated by the user's relevance assessments. The reformulation can be automatic (e.g., automatic relevance feedback) or manual (e.g., the user selects additional or alternative terms to reformulate the query). The process can be viewed as a flow of information between the different steps since the information in a step comes from the information contained in previous steps.

There are successive steps in a query reformulation process; these are represented by the sequential composition of channels. There can also be parallel methods used to reformulate a query; these can be represented with the parallel combination of channels. The constraints defining the nature of the transitions depend on how the reformulation process is implemented. In [16], Koenemann and Belkin discuss three approaches to query reformulation: *opaque* in which the users are not shown the terms to be added to the query, *transparent* in which the users are shown the terms that had been added to the query after the new query had been executed, and the *penetrable* case in which the system allows the user to select which terms to add to the query before execution. In addition, the user can manually add or remove query terms in all cases. We shall consider these cases in the following example.

The following steps occur in all cases after the initial query and display of retrieved documents to the user:

- **Step 1:** The user marks the relevant documents, and possibly the irrelevant documents.
- **Step 2:** The system generates a list of query expansion terms.
- **Step 3:** The user modifies the original query (by adding or removing terms).
- **Step 4:** Query expansion (the variable condition in Koenemann and Belkin's experiments) is performed.
- **Step 5:** The modified query is submitted to the system.
- **Step 6:** Documents are retrieved and presented to the user.
- **Step 7:** Based on the new set of documents the user decides whether to continue.

Either the user continues with the relevance feedback (the user continues to mark relevant documents, returning to step one), or the process finishes (for example the user has enough relevant information, the performance has decreased to the point that the use does not want to continue modifying the query). We model the interaction up to step 7 only.

Let us assume that some data have been collecting about how a number of users have reformulated their queries. Suppose that all the 7 above steps where indeed observed. We represent each step i by a situation s_i. For the opaque and transparent cases the information supported by the situations are shown in the following table:

decision is the decision made whether to stop the query transformation process, or whether to continue it. In the penetrable case, the user chooses which system generated terms to add. We assert a new type $USAT$ to describe the terms the user has selected from the system generated list (reprsented by the type T). We will have then $s_4 \models USAT$, instead of $s_4 \models SAT$.

Suppose that 7 (single) transitions have been identified. These are represented by the 7 channels below. The situations linked by and the constraints supported by each channel for the three cases are shown in the following table:

Situations	Types
s_1	query Q, retrieved relevant documents D_{ret_rel} retrieved irrelevant documents D_{ret_irrel}
s_2	expansion terms T
s_3	query Q, user added terms UAT, user removed terms URT
s_4	user transformed query Q', system added terms SAT
s_5	transformed query Q''
s_6	transformed query Q'', retrieved documents D'_{ret}
s_7	*decision*

Table 1: Query formulation: situations

Opaque/Transparent				Penetrable			
c_1	s_1	s_2	$Q \wedge D_{ret\_rel} \wedge D_{ret\_irrel} \rightarrow T \wedge Q$				
c_2	s_2	s_4	$T \rightarrow SAT$	c_2	s_2	s_4	$T \rightarrow USAT$
c_3	s_2	s_3	$Q \rightarrow UAT \wedge URT \wedge Q$				
c_4	s_3	s_4	$Q \wedge UAT \wedge URT \rightarrow Q'$				
				c_4^{-1}	s_4	s_3	$Q' \rightarrow UAT \wedge URT \wedge Q'$
c_5	s_4	s_5	$SAT \wedge Q' \rightarrow Q''$	c_5	s_4	s_5	$USAT \wedge Q' \rightarrow Q''$
c_6	s_5	s_6	$Q'' \rightarrow D_{ret'}$				
c_7	s_5	s_6	$D_{ret'} \rightarrow decision$				

Table 2: Query formulation: channels

We discuss channel c_4. In all cases, from the query Q, users added terms UAT and users removed terms URT, the query is reformulated (manually) Q'. In the opaque and transparent cases, when in situation s_4, users cannot go back to s_3 (adding or removing query terms). In the penetrable case, as the user chooses which system generated terms to add, user are allowed to move between s_3 and s_4. That is, the they can move freely between adding their own terms and adding terms selected by the system. This is represented by having the inverse channel c_4^{-1}, in addition to channel c_4. The constraint supported by the inverse channel is $Q' \rightarrow UAT \wedge URT \wedge Q'$, since the situation s_3 will support now the reformulated query Q' and the user added/removed terms with respect to Q'.

In the opaque/transparent case, the interaction is represented by the following sequence of channels: $c_1;(c_2||(c_3;c_4);c_5;c_6$. The information supported by the situation s_4 comes from two parallel channels, c_2 yielding the system added terms SAT, and $c_3; c_4$ yielding the user reformulated query Q'. The latter involves an intermediary step, represented by situation s_3, where the users' added and removed terms are determined.

A system that does not allow a user to modify the query but only allows automatic query expansion would have a structure similar to that of the opaque/transparent case, but without the manual query modification process, which is represented by the sequence of channel $c_3; c_4$. The interaction is then modelled by the following sequence of channels: $c_1;c_2;c_5;c_6$.

5.2 Web search

We model the standard interaction between a user (such as ourselves) and a search engine such as AltaVista (we do not consider the case when a user is browsing various advertising displayed by the web search engine). For a typical information seeking task, the steps are as follows:

- **Step 1:** The user formulates the query.
- **Step 2:** (10) Documents are retrieved and displayed to the user.
- **Step 3:** Based on the retrieved documents, the user decides what to do next. If the user is interested by the result, then he/she will browse the result (Step 4). Otherwise, the user will ask for further documents to be retrieved (Step 2), or will formulate a new query (Step 1). Alternatively, the user could either refine the query (Step 5) or stop the search (Step 6).

- **Step 4:** The user browses the result.
- **Step 5:** The user defines the refinement parameters.
- **Step 6:** The user stops the search.

Each step i is modelled by a situation s_i. The types (information) supported by the situations are shown in the following table:

Situations	Types
s_1	query Q
s_2	query Q, retrieved documents D_{ret}
s_3	query Q, retrieved documents D_{ret}, $decision$
s_4	retrieved documents D_{ret}, $browsing$
s_5	query Q, refinement parameters ref_{par}
s_6	$stop$

Table 3: Web search: situations

$decision$ is the type of situations where a user is making a decision regarding the next step. $browse$ is the type of situations where a user is browsing the result (looking at the list of displayed documents or their summary, or browsing the documents themselves). $stop$ is the type of situations where a user has stopped his/her information seeking activity.

The transitions between situations are modelled by channels. To model a standard search, we use 7 channels $c_1, \ldots, c_7$. The source and target situations and the constraints supported by each channel are shown in the following table:

Channels	Sources	Targets	Constraints
c_1	s_1	s_2	$Q \rightarrow Q \wedge D_{ret}$
	s_3	s_2	$Q \rightarrow D_{ret} \vert further$
c_2	s_2	s_3	$D_{ret} \wedge Q \rightarrow D_{ret} \wedge Q \wedge decision$
	s_4	s_3	$D_{ret} \rightarrow decision$
c_3	s_3	s_4	$D_{ret} \rightarrow browse \vert interest$
c_4	s_3	s_5	$D_{ret} \wedge Q \rightarrow ref_{par} \wedge Q \vert refine$
c_5	s_5	s_1	$ref_{par} \wedge Q \rightarrow Q' \vert satisfied$
c_6	s_3	s_6	$D_{ret} \rightarrow stop \vert satisfied$
c_7	s_3	s_6	$D_{ret} \rightarrow stop \vert enough$
1	s_5	s_5	$D_{ret} \wedge Q \wedge ref_{par} \rightarrow ref'_{par} \vert refine$
	s_4	s_4	$D_{ret} \rightarrow browse \vert interest$

Table 4: Web search: channels

$further$ is the type of situations where a user wants to see further retrieved documents. $interest$ is the type of situations where a user is interested by the result (for example, the user states that he/she is finding some documents that seem relevant to his/her information need). $refine$ is the type of situations where a user wants to refine the query. For this, the user must set some parameters which are represented by the type ref_{par}. $satisfied$ is the type of situations where a user states his/her satisfaction(e.g., about the result, the refinement parameters). Finally, $enough$ is the type of situations where a user expresses that he/she does not want to use the system any longer (e.g., the relevant documents have been found, the user is frustrated with the system, or the user has no more time).

Let us take for example channel c_3 linking s_3 to s_4. The nature of the channel is $D_{ret} \rightarrow browse|interest$. Information flows from s_3 to s_4 because the user is browsing the retrieved documents. The transition is uncertain, which is represented by the background conditions $interest$; it depends on whether the user is interested in the set of retrieved documents, that is $s_3 \models interest$.

We have two instances of unity channel: the user continues browsing the result ($s_4 \models D_{ret} \xmapsto{1} s_4 \models browse$) and the user refines the query further ($s_5 \models D_{ret} \wedge Q \wedge ref_{par} \xmapsto{1} s_5 \models ref'_{par}$). Note that in both cases, we have uncertainty. For the transitions to occur, it must be the case that, respectively, $s_4 \models interest$ and $s_5 \models refine$. ref'_{par} represents the new set of refinement parameters.

Now we discuss a reasoning example. Consider the following sequence of channels: $C_1 = c_1; c_2; c_3; c_2; c_6$. This models that a user is satisfied after browsing the first set of retrieved documents. If one observes that the interaction of most users can be modelled by the sequential channel C_1, then one can derive that the other functionalities offer by the search engine are not used. Consider now the following sequence of channels: $C_2 = c_1; c_2; c_3; c_2; c_4; 1; 1; 1; c_5; c_1; c_2; c_6$. The sequences of transitions modelled by C_1 and C_2 lead to the satisfaction of the user (the transition modelled by channel c_6 occurs in both cases when the user is satisfied), but C_2 involves an iterative query refinement (represented by the three occurrences of the unity channel $s_5 \stackrel{1}{\longmapsto} s_5$). Suppose that one observes for a given a user that the documents retrieved without refinement contain those retrieved with refinement. This can be represented as follows. Let $C'_1 = c_1; c_2$ and $C'_2 = c_1; c_2; c_3; c_2; c_4; 1; 1; 1; c_5; c_1; c_2$. The source and target situations for the two channels are s_1 and s_2, respectively. We have then $s_1 \models Q$ and $s_2 \models D_{ret}$. Therefore, the information obtained using C'_1 and C'_2 in parallel could be obtained using C'_1 alone. This is represented as $C'_1 || C'_2 = C'_1$, that is $C'_1 \preceq C'_2$. What can be inferred is that the query refinement process is ineffective for the given user.

6 Conclusion

This paper presents a formal framework based on channel theory for modelling the interaction between a user and an IR system. The aim of such a framework is to obtain a better understanding of the nature of the interaction in IR, thus leading to more effective retrieval.

Channel theory as a formal framework for modelling interactive IR systems has the potential to be a very powerful tool. It is capable of encapsulating both system and user behaviour, from very specific individual actions to high level conceptual goals. It can not only model the information contained within documents but also the information given by the interaction itself. Finally, as it is a domain- and system-independent representation language, its power is not restricted to only describing the interaction; channel theory allows us to predict and reason about the interaction.

We have applied the framework to model interaction in relevance feedback and standard web search. What the examples show is by formally modelling the interaction, we can derive facts about the interaction between users and IR systems. These facts can then be used to enhance the support facilities offered by the system. A major advantage is that we can produce set of facts describing the interaction *within* a system, or *across* systems. With other models of interaction, this was not possible.

We see two directions to follow from this work. The first one is to apply our framework to model interaction in actual information seeking cases (real users with real IR systems), with the aim to derive principles that would allow for more effective use of IR systems. The second direction is to relate our model to those that have been developed. These models present strong study of the interaction in IR. These studies could be mapped to our framework, thus allowing reasoning specific to the scope (e.g., complexity of the search, categorisation of search) addressed by these models.

References

[1] J. Barwise. *The Situation in Logic*. CLSI Lectures Notes 17, Stanford, California, 1989.

[2] J. Barwise. Constraints, channels and the flow of information. In *Situation Theory and its Application*, volume III, Stanford, California, 1993. CLSI Lectures Notes 37.

[3] M. Bates. Where should the person stop and the information search interface start? *Information, Processing and Management*, 26(5):575–591, 1990.

[4] M. Beaulieu and S. Jones. Interactive searching and interface issues in the okapi best match probabilistic retrieval system. *Interacting with computers*, 10(3):237 –248, 1998.

[5] N.J. Belkin, C. Cool, A. Stein, and U. Thiel. Cases, scripts and information-seeking strategies: On the design of interactive information retrieval systems. *Expert Systems with Applications*, 29(3):325–344, 1993.

[6] S. Park C. Cool, N. J. Belkin, J. Koenemann, and K.B. Ng. Information seeking behavior in new searching environment. In *CoLIS 2, Second International Conference on Conceptions of Library and Information Science: Integration in Perspective*, pages 403–416, Copenhagen, Royal School of Librarianship, 1996.

[7] J.P. Callan, W.B. Croft, and S.M. Harding. The inquery retrieval system. In *International Conference on Database and Expert Systems Applications*, pages 78–83, 1992.

[8] K. J. Devlin. *Logic and Information*. Cambridge University Press, Cambridge, England, 1991.

[9] D. Ellis. A behavioural approach to informatino system design. *Journal of Documentation*, 45(3):171–212, 1989.

[10] D. Ellis, D. Cox, and K. Hall. A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, 49:356–369, 1993.

[11] V. Florance and G. Marchionini. Information processing in the context of medical care. In *ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 158–163, Zurich, Switzerland, 1996.

[12] N. Ford and R. Ford. Towards a cognitive theory of information accessing: an empirical study. *Information Processing and Managemen*, 29(5):569–585, 1993.

[13] D. Harman, editor. *Overview of the Fourth Text Retrieval Conference*. NIST Special Publication, 1996.

[14] T. W. C. Huibers, M. Lalmas, and C. J. van Rijsbergen. Information Retrieval and Situation Theory. *SIGIR Forum*, 30(1):11–25, 1996.

[15] P. Ingwersen. *Information retrieval interaction*. Taylor Graham, 1992.

[16] J. Koenemann and N.J. Belkin. A case for interaction: a study of interactive information retrieval behaviour and effectiveness. In *CHI'96*, 1996.

[17] C.C. Kuhlthau. Inside the search process: information seeking from the user's perspective. *Journal of the American Society of Information Science*, 42:361–371, 1991.

[18] C.C. Kuhlthau. Principle for uncertainty for information seeking. *Journal of Documentation*, 49(4):339–355, 1993.

[19] M. Lalmas. The flow of information in information retrieval: Towards a general framework for the modelling of information retrieval. In *Information Retrieval: Uncertainty and Logics - Advanced models for the representation and retrieval of information*, pages 129–150. Kluwer Academic Publishers, 1998.

[20] M. Magennis and C. J. van Rijsbergen. The potential and actual effectiveness of interactive query expansion. In *ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 324–332, Philadelphia, USA, 1997.

[21] S.E. Robertson and K. Sparck Jones. Relevance weighting of search terms. *Journal of the American Society of Information Science*, 27(5):129–146, 1976.

[22] C. J. van Rijsbergen and M. Lalmas. An Information Calculus for Information Retrieval. *Journal of the American Society of Information Science*, 47(5):385–398, 1996.

[23] C.J. van Rijsbergen, D. Harper, and M. Porter. The selection of good search terms. *Information, Processing and Management*, 17:77–91, 1981.

[24] E.M. Voorhees. Query expansion using lexical semantic relations. In *ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 61–69, Dublin, Ireland, 1994.